H-NS Silences *gfp*, the Green Fluorescent Protein Gene: *gfp*^{TCD} Is a Genetically Remastered *gfp* Gene with Reduced Susceptibility to H-NS-Mediated Transcription Silencing and with Enhanced Translation$^\nabla$§

Colin P. Corcoran, Andrew D. S. Cameron, and Charles J. Dorman*

Department of Microbiology, Moyne Institute of Preventive Medicine, School of Genetics and Microbiology, Trinity College Dublin, Dublin 2, Ireland

The bacterial nucleoid-associated protein H-NS, which preferentially targets and silences A+T-rich genes, binds the ubiquitous reporter gene *gfp* and dramatically reduces local transcription. We have redesigned *gfp* to reduce H-NS-mediated transcription silencing and simultaneously improve translation *in vivo* without altering the amino acid sequence of the GFP protein.

The green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* is widely used as a reporter of transcriptional activity and protein localization in both eukaryotes and prokaryotes (reviewed in reference 16). The DNA encoding *gfp* and many of its variants has a high A+T content (~60%) and strong predicted DNA curvature, both of which are key features of DNA that is targeted by the histone-like nucleoid structuring protein (H-NS) (10, 12, 14, 15, 22). H-NS is a global gene silencer that represses transcription by condensing chromosome and plasmid DNA in many Gram-negative bacteria. For example, H-NS binds and represses ~15% of the *Escherichia coli* and *Salmonella enterica* genomes (13), including A+T-rich genes that have been acquired horizontally (5). This raised the possiblity that H-NS may bind and repress the alien *gfp* gene and thus obscure measurements of promoter activity.

Many variants of GFP have been developed to optimize this versatile reporter for different applications. The alteration of these major characteristics of the protein have typically been achieved by making a very small number of amino acid substitutions involving only minor alterations to the DNA sequence of *gfp*. For example, the alteration of just one amino acid, arising from a single nucleotide substitution, was required to convert GFP to blue fluorescent protein (BFP) (8). Thus, many of the commonly used variants of GFP are encoded by genes that are ideal targets for H-NS and H-NS-like proteins.

The widely used and highly fluorescent *gfp+* (17) has an A+T content of 59% and was predicted by the Bend.it algorithm (21) to contain several regions of curved DNA (Fig. 1A), making it a likely target for H-NS binding. The *gfp+* gene was therefore redesigned to have an A+T content that was similar to that of the *E. coli* chromosome (49%). Gene Designer (20) was used both to reduce the A+T content of *gfp+* and to optimize its codon usage for expression in *E. coli*. This approach produced iterative, equally optimized sequences that were screened for reduced intrinsic DNA curvature using Bend.it. The preferred sequence differed from *gfp+* at 157 positions of the 717-base-pair gene, had 50% A+T content, and showed dramatically reduced predicted DNA curvature (Fig. 1A). This reengineered gene is referred to as *gfp*^{TCD} (TCD for Trinity College Dublin), and its DNA sequence is provided in Fig. S1 in the supplemental material.

Electrophoretic mobility shift assays (EMSA) were performed to determine the affinity of H-NS for *gfp+* and *gfp*^{TCD} *in vitro*. H-NS bound to *gfp+* with an apparent dissociation constant (K_d) of 4.9 ± 0.1 nM (Fig. 1B). A narrow protein concentration range (4.5 to 10.55 nM) was required for the transition from initial binding to the fully bound DNA probe that resolved as a single low-mobility complex. In contrast, H-NS displayed a relatively lower affinity for *gfp*^{TCD} (K_d, 7.5 ± 0.5 nM) (Fig. 1B), resulting in the formation of a poorly resolved H-NS–DNA complex over a wider range of protein concentrations (7.9 to 18.75 nM); the *gfp*^{TCD} probe resolved as a single H-NS–DNA complex only at protein concentrations above 25 nM.

The *E. coli proU* transcriptional regulatory region was used as a positive control for H-NS binding because it contains several well-characterized H-NS binding sites (1). As expected, the *proU* DNA probe was bound by H-NS (K_d, 6.2 ± 0.5 nM) in EMSA where it resolved as two separate H-NS–DNA complexes (Fig. 1C). H-NS has a relatively lower binding affinity for the *E. coli lacZ* gene (15), so this DNA was used as a negative control; EMSA analysis confirmed that *lacZ* was indeed more weakly bound by H-NS (K_d, 16.3 ± 1.8 nM) than was *proU* (Fig. 1C).

The *proU* operon was also used to test the effect of H-NS binding to *gfp+* on transcription from a nearby promoter. Transcriptional repression of the *proU* operon in low-osmolarity conditions involves binding by H-NS to the A+T-rich, highly curved DNA located in the upstream and downstream regulatory elements (URE and DRE, respectively) that flank the *proU* promoter (1, 11, 15). The repressive complex formed
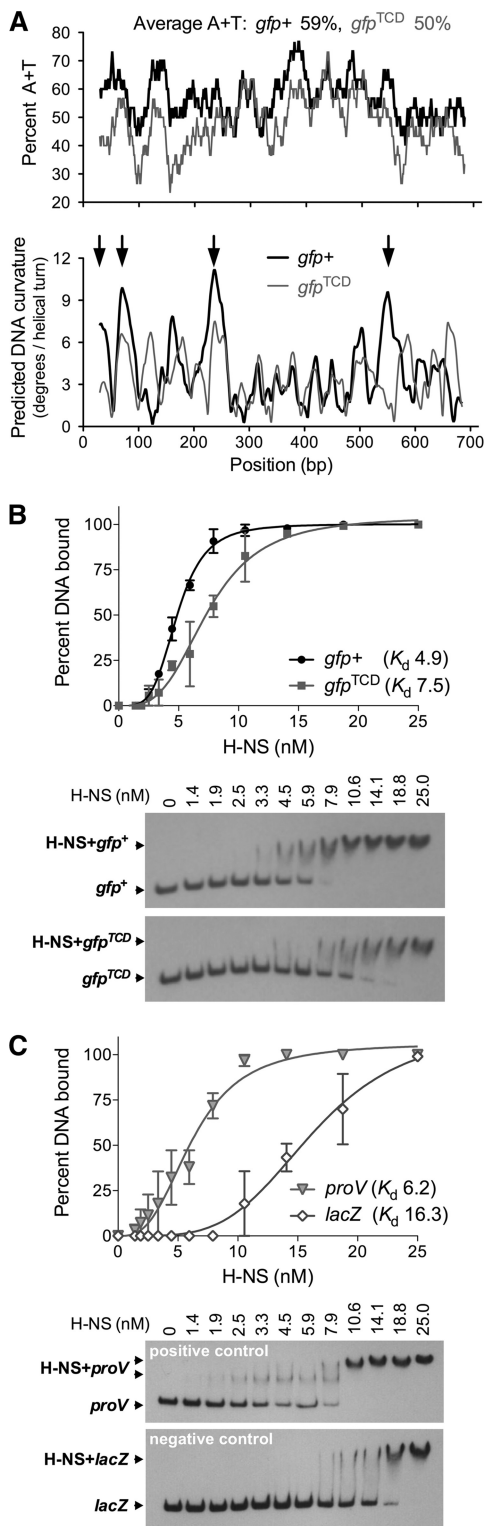
FIG. 1. DNA sequence features of the *gfp+* gene are conducive to H-NS binding. (A) A+T content (upper plot) and predicted intrinsic DNA curvature (lower plot) of *gfp+* and *gfp*^TCD. Arrows highlight regions of strong predicted curvature in *gfp+* that are reduced in *gfp*^TCD. (B and C) EMSA analysis of H-NS binding to *gfp+* and *gfp*^TCD (B) and *proU* and *lacZ* (C). Binding curves and K_d values are shown in the upper plots; mean values and ranges are plotted. Typical EMSA gels from which binding curves were derived are shown below. See supplemental material for a complete description of the experimental methods.
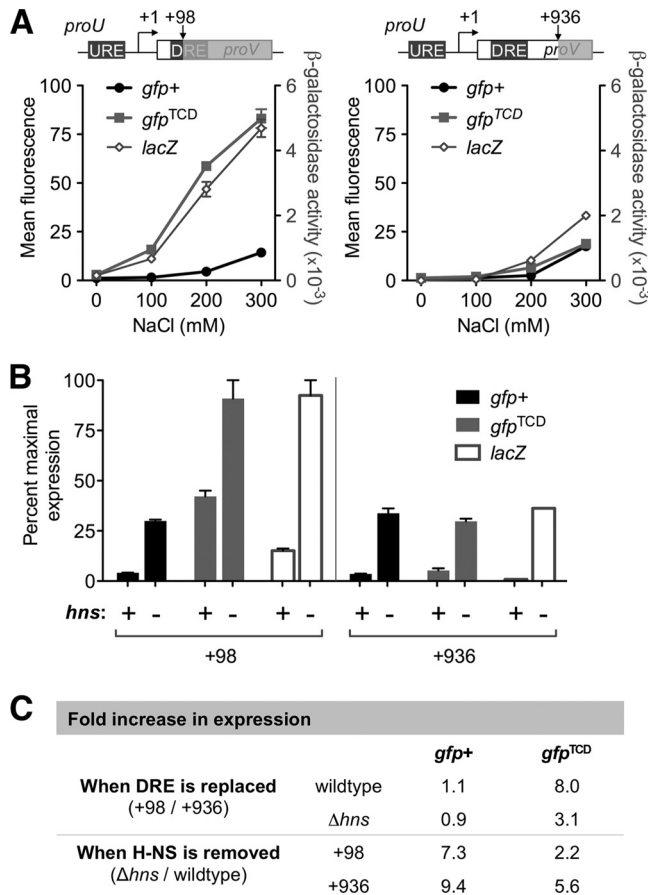

FIG. 2. *gfp+* and *gfp*^TCD have different effects on transcription and translation. (A) The effect of increasing NaCl concentrations on the expression of reporter gene fusions inserted at +98 and +936 bp from the *proU* transcriptional start site. Mean values and ranges are plotted. Diagrams not to scale. (B) The effect of removing H-NS on fusion gene expression in 100 mM NaCl. Mean values and ranges are plotted. (C) Fold increases in fusion gene expression caused by changes in H-NS occupancy. See supplemental material for a complete description of the experimental methods.

by H-NS at *proU* is disrupted by increasing the osmolarity of the medium (7). While H-NS binding in the DRE is essential for osmoregulation of *proU*, it has previously been shown that the native DRE can be functionally replaced by an unrelated piece of DNA that is bound by H-NS (15). It was also shown that the *lacZ* gene cannot functionally replace the DRE and thus its presence in place of the DRE leads to derepression of *proU* in low-osmolarity conditions (100 mM NaCl) (15). Therefore, we tested the ability of *gfp+*, *gfp*^TCD, or *lacZ* to functionally replace the DRE by insertion of each gene within the DRE at +98 bp downstream from the transcriptional start site (Fig. 2A). Expression of the *proU-gfp+*(+98) fusion was repressed relative to expression of *proU-gfp*^TCD(+98), indicating that *gfp+*, but not *gfp*^TCD, functionally replaced the DRE as a binding site for H-NS and thus repressed *proU* expression. The *proU-lacZ*(+98) fusion was not repressed at low NaCl concentrations, consistent with earlier observations (15). Reporter fusions that did not disrupt the DRE were generated by insertion of each of the three reporter genes (*gfp+*, *gfp*^TCD, and

lacZ) at +936 bp downstream from the transcriptional start site (Fig. 2A). As predicted, expression of all three *proU*(+936) fusions was repressed by H-NS binding to the DRE at NaCl concentrations below 200 mM.

To confirm that H-NS binding to *gfp*+ accounted for repression of *proU-gfp*+(+98) *in vivo*, the expression of all *proU* fusions was tested in a Δ*hns* background. Cells were cultured in the repressive conditions of 100 mM NaCl, and the data were expressed as a percentage of maximal derepressed expression to facilitate comparisons between GFP fluorescence and β-galactosidase activity (Fig. 2B). For all three fusions, *proU* expression was elevated in the absence of H-NS. This revealed that even in the absence of the DRE, H-NS continued to bind the URE and repress the *proU-gfp*^TCD(+98) and *proU-lacZ*(+98) fusions. These data allowed an assessment of the relative effects of replacing only the DRE compared to those of eliminating H-NS protein from the cell (Fig. 2C). Replacing the DRE with *gfp*^TCD resulted in an 8-fold increase in expression relative to expression from the *proU-gfp*^TCD(+936) fusion; a similar comparison between *proU-gfp*+ at positions +98 and +936 showed that *gfp*+ inserted in the DRE maintained full repression. Consequently, eliminating H-NS genetically resulted in a 7.3- to 9.4-fold increase in expression from the *gfp*+ fusions while the expression of *proU-gfp*^TCD(+98) improved only 2.2-fold upon removal of H-NS (Fig. 2C). This also confirmed that H-NS-mediated repression of *proU* occurs primarily through binding to the DRE (Fig. 2C).

Even in the absence of H-NS, *proU-gfp*^TCD(+98) produced 3-fold-higher levels of GFP than *proU-gfp*+(+98) (Fig. 2C). The *gfp*^TCD gene is codon optimized for translation in *E. coli*, and these data suggest that *gfp*^TCD is translated 3-fold more efficiently than *gfp*+. Improved translation efficiency was confirmed by cloning *gfp*+ and *gfp*^TCD in the pPro vector under the control of the propionate-inducible *prpBCDE* promoter (9), for which there is no evidence of H-NS binding or repression (3). In pPro, *gfp*^TCD produced on average 3.5-fold more GFP in a wild-type background and 2.7-fold more GFP in a Δ*hns* background than *gfp*+ (Fig. 3A). The codon adaptation index (CAI) is a standard means to calculate the effects of species-specific codon biases on translation (18); Fig. 3B shows that *gfp*^TCD is predicted to have improved translation efficiency in both bacterial and eukaryotic model organisms. *gfp*^TCD therefore provides an ideal template for the directed evolution of novel *gfp* variants.

In many bacteria, including *E. coli*, H-NS selectively targets horizontally acquired genes of high A+T content preventing transcription of the potentially deleterious foreign DNA (5). The introduction of A+T-rich plasmid DNA can thus titrate H-NS away from native chromosomal locations, leading to a mild Δ*hns* phenotype with pleiotropic effects, including reduced virulence and motility (3, 6). We have shown that H-NS binds strongly in the *gfp*+ gene, perhaps explaining a recent report of reduced invasiveness of *Salmonella* due solely to the presence of a promoterless *gfp*+ gene on a multicopy plasmid (2). Over the years, many bacterial studies have relied on plasmid-based *gfp* fusions, and some of these studies may have experienced unrecognized pleiotropic effects due to H-NS titration. The *gfp* gene has been used in conjunction with plasmid-based high-throughput genetic screens in many landmark studies, including the identification of promoters involved in
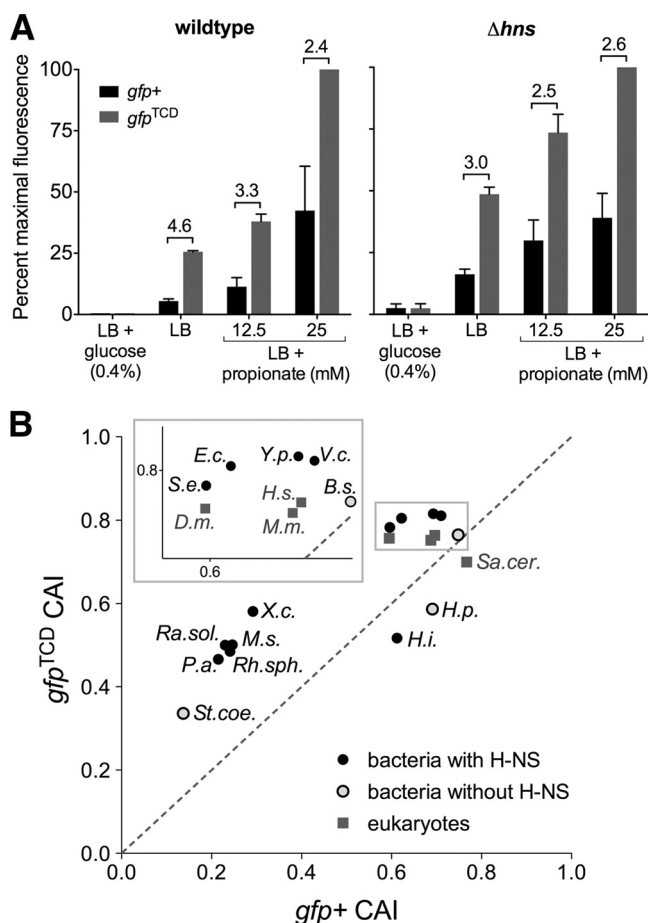

FIG. 3. *gfp*^TCD has improved translation efficiency. (A) Expression of *gfp*+ and *gfp*^TCD from the *prpBCDE* promoter on the plasmid pPro in wild-type and Δ*hns* cells. The *prpBCDE* promoter is repressed by glucose and induced by propionate. Mean values and ranges are plotted; fold differences between *gfp*+ and *gfp*^TCD fluorescence levels are shown above the bars. (B) Codon adaptation index (CAI) values for *gfp*^TCD and *gfp*+ in various organisms. Organisms listed above the dashed line will translate *gfp*^TCD more efficiently than *gfp*+ (and wild-type *gfp*). Organism classification and abbreviations are as follows. *Alphaproteobacteria*: *Rh.sph.*, *Rhodobacter sphaeroides*. *Betaproteobacteria*: *Ra.sol.*, *Ralstonia solanacearum*. *Epsilonproteobacteria*: *H.p.*, *Helicobacter pylori*. *Gammaproteobacteria*: *E.c.*, *Escherichia coli*; *H.i.*, *Haemophilus influenzae*; *P.a.*, *Pseudomonas aeruginosa*; *S.e.*, *Salmonella enterica*; *V.c.*, *Vibrio cholerae*; *X.c.*, *Xanthomonas campestris*; *Y.p.*, *Yersinia pestis*. *Actinobacteria*: *M.s.*, *Mycobacterium smegmatis*; *St.coe.*, *Streptomyces coelicolor*. *Firmicutes*: *B.s.*, *Bacillus subtilis*. *Eukaryota*: *D.m.*, *Drosophila melanogaster*; *H.s.*, *Homo sapiens*; *M.m.*, *Mus musculus*; *Sa.cer.*, *Saccharomyces cerevisiae*. See supplemental material for a complete description of the experimental methods.

the intracellular survival of *Salmonella enterica* serovar Typhimurium (19) and the adaptation by *E. coli* to metabolisms of different carbon sources (23).

Since nucleoid-associated proteins with a preference for A+T-rich DNA are found in most cell types, including archaea and eukaryotes (4), our discovery of H-NS binding and repression of *gfp* may have far-reaching implications for the use of *gfp* in all domains of life. The development of *gfp*^TCD, which is both a poor target for H-NS and a more accurate reporter of transcription that avoids H-NS-mediated downregulation of

the promoter under investigation, provides a new tool for future studies free from artifacts caused by H-NS interference.

REFERENCES

1. **Bouffartigues, E., M. Buckle, C. Badaut, A. Travers, and S. Rimsky.** 2007. H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. Nat. Struct. Mol. Biol. **14:**441–448.
2. **Clark, L., I. Martinez-Argudo, T. J. Humphrey, and M. A. Jepson.** 2009. GFP plasmid-induced defects in *Salmonella* invasion depend on plasmid architecture, not protein expression. Microbiology **155:**461–467.
3. **Dillon, S. C., A. D. S. Cameron, K. Hokamp, S. Lucchini, J. C. Hinton, and C. J. Dorman.** 2010. Genome-wide analysis of the H-NS and Sfh regulatory networks in *Salmonella* Typhimurium identifies a plasmid-encoded transcription silencing mechanism. Mol. Microbiol. **76:**1250–1265.
4. **Dillon, S. C., and C. J. Dorman.** 2010. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. Nat. Rev. Microbiol. **8:**185–195.
5. **Dorman, C. J.** 2007. H-NS, the genome sentinel. Nat. Rev. Microbiol. **5:**157–161.
6. **Doyle, M., M. Fookes, A. Ivens, M. W. Mangan, J. Wain, and C. J. Dorman.** 2007. An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. Science **315:**251–252.
7. **Gowrishankar, J.** 1985. Identification of osmoresponsive genes in *Escherichia coli*: evidence for participation of potassium and proline transport systems in osmoregulation. J. Bacteriol. **164:**434–445.
8. **Heim, R., D. C. Prasher, and R. Y. Tsien.** 1994. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. Proc. Natl. Acad. Sci. U. S. A. **91:**12501–12504.
9. **Lee, S. K., and J. D. Keasling.** 2005. A propionate-inducible expression system for enteric bacteria. Appl. Environ. Microbiol. **71:**6856–6862.
10. **Lucchini, S., G. Rowley, M. D. Goldberg, D. Hurd, M. Harrison, and J. C. Hinton.** 2006. H-NS mediates the silencing of laterally acquired genes in bacteria. PLoS Pathog. **2:**e81.
11. **Lucht, J. M., P. Dersch, B. Kempf, and E. Bremer.** 1994. Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled proU operon of *Escherichia coli*. J. Biol. Chem. **269:**6578–6586.
12. **Navarre, W. W., S. Porwollik, Y. Wang, M. McClelland, H. Rosen, S. J. Libby, and F. C. Fang.** 2006. Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. Science **313:**236–238.
13. **Noom, M. C., W. W. Navarre, T. Oshima, G. J. Wuite, and R. T. Dame.** 2007. H-NS promotes looped domain formation in the bacterial chromosome. Curr. Biol. **17:**R913–914.
14. **Oshima, T., S. Ishikawa, K. Kurokawa, H. Aiba, and N. Ogasawara.** 2006. *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. DNA Res. **13:**141–553.
15. **Owen-Hughes, T. A., G. D. Pavitt, D. S. Santos, J. M. Sidebotham, C. S. Hulton, J. C. Hinton, and C. F. Higgins.** 1992. The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. Cell **71:**255–265.
16. **Rizzo, M. A., M. W. Davidson, and D. W. Piston.** 1 December 2009, posting date. Fluorescent protein tracking and detection: fluorescent protein structure and color variants. Cold Spring Harb. Protoc. doi:10.1101/pdb.top63.
17. **Scholz, O., A. Thiel, W. Hillen, and M. Niederweis.** 2000. Quantitative analysis of gene expression with an improved green fluorescent protein. Eur. J. Biochem. **267:**1565–1570.
18. **Sharp, P. M., and W. H. Li.** 1987. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. **15:**1281–1295.
19. **Valdivia, R. H., and S. Falkow.** 1997. Fluorescence-based isolation of bacterial genes expressed within host cells. Science **277:**2007–2011.
20. **Villalobos, A., J. E. Ness, C. Gustafsson, J. Minshull, and S. Govindarajan.** 2006. Gene Designer: a synthetic biology tool for constructing artificial DNA segments. BMC Bioinformatics **7:**285.
21. **Vlahovicek, K., L. Kajan, and S. Pongor.** 2003. DNA analysis servers: plot.it, bend.it, model.it and IS. Nucleic Acids Res. **31:**3686–3687.
22. **Yamada, H., S. Muramatsu, and T. Mizuno.** 1990. An *Escherichia coli* protein that preferentially binds to sharply curved DNA. J. Biochem. **108:**420–425.
23. **Zaslaver, A., A. Bren, M. Ronen, S. Itzkovitz, I. Kikoin, S. Shavit, W. Liebermeister, M. G. Surette, and U. Alon.** 2006. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. Nat. Methods **3:**623–628.